Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 11513–11517, December 1992
Biophysics

Poly(L-alanine) as a universal reference material for understanding protein energies and structures

TERESA HEAD-GORDON, FRANK H. STILLINGER, MARGARET H. WRIGHT, AND DAVID M. GAY

AT&T Bell Laboratories, Murray Hill, NJ 07974

Contributed by Frank H. Stillinger, June 17, 1992

ABSTRACT **We present a proposition, the "poly(L-alanine) hypothesis," which asserts that the native backbone geometry for any polypeptide or protein of *M* residues has a closely mimicking, mechanically stable, image in poly(L-alanine) of the same number of residues. Using a molecular mechanics force field to represent the relevant potential energy hypersurfaces, we have carried out calculations over a wide range of *M* values to show that poly(L-alanine) possesses the structural versatility necessary to satisfy the proposition. These include poly(L-alanine) representatives of minima corresponding to secondary and supersecondary structures, as well as poly(L-alanine) images for tertiary structures of the naturally occurring proteins bovine pancreatic trypsin inhibitor, crambin, ribonuclease A, and superoxide dismutase. The successful validation of the hypothesis presented in this paper indicates that poly(L-alanine) will serve as a good reference material in thermodynamic perturbation theory and calculations aimed at evaluating relative free energies for competing candidate tertiary structures in real polypeptides and proteins.**

Section 1: Introduction

In this paper we present an idea, the "poly(L-alanine) hypothesis," for the purpose of offering an independent approach to understanding protein structure and energetics. It asserts that there is a close correspondence, or short-distance mapping, between the backbone geometry at the native-structure free-energy minimum for any polypeptide or protein of *M* residues and that of a (metastable) minimum of poly(L-alanine) of the same number of residues. Fig. 1 illustrates this concept. Alanine is the most natural choice of residue for such a comprehensive comparison because its degree of geometric plasticity, as measured by a Ramachandran plot (1), is most representative of all amino acids except glycine and proline. Furthermore it is the simplest chiral residue. Notice that inverse correspondences are *not* postulated: for steric reasons the global free energy minimum for $(\text{L-Ala})_M$ may have no close image among the local minima for the free-energy function of a given *M*-residue protein.

Several conceptual advantages arise from successful validation of the poly(L-alanine) hypothesis. One is that *mechanical* stability of protein native structures do not depend on side-chain details but that *absolute* energy (or free energy) stability is indeed controlled by those details. In this connection Matthews and coworkers' (2, 3) T4-lysozyme mutagenesis studies have a special significance, wherein replacement of several residues by L-alanine has been demonstrated to preserve native structure. A second advantage is that poly(L-alanine) can serve as a natural reference material in theoretical calculations of the relative stabilities of distinct folded structures for a given protein or of the relative stabilities of distinct *M*-residue proteins in the same tertiary structure. A third advantage, as explained in *Section 4*, is that the hypothesis provides a context in which the specific structural roles of disulfide bonds, charged side chains, packing of amino acid side chains, and special residues such as proline and glycine can be independently and quantitatively assessed.

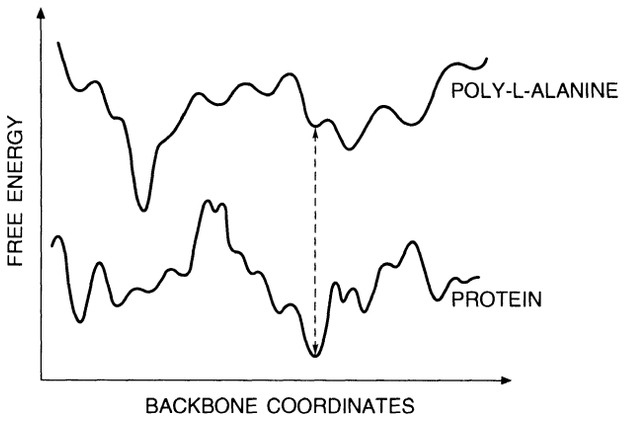

FIG. 1. Topographic basis of the poly(L-alanine) hypothesis. Free energy hypersurfaces are compared for a given protein and for poly(L-alanine) with the same number of residues.

A variety of calculations has been undertaken to demonstrate the validity of the poly(L-alanine) hypothesis. These calculations range in *M*, the number of residues, from very small to quite large polypeptides and demonstrate the ability of poly(L-alanine) to adopt an impressive repertoire of structural alternatives. This diversity is evident even at the dimer level (4). Using larger oligomers, we have demonstrated stability for representative secondary structures (helices, turns, sheets), for supersecondary conformers, and for tertiary structures exhibited by several naturally occurring proteins.

Section 2: Methods

Potential Energy Functions. The empirical potential energy function used has the typical molecular mechanics form and has been described elsewhere (4). We have used the parameters of both the all-atom representation (5) and the extended-atom representation (version 19) of CHARMM (6). For reasons of expediency, the all-atom representation (5) has been used for the cases where the polypeptide has 10 amino acids or fewer; in all other cases the extended-atom representation was used (6). For the smaller cases we have also used methyl blocking groups instead of the amino- and carboxyl-terminating groups of naturally occurring peptides. For the larger cases we have used the zwitterionic form for initiating and terminating the polypeptide.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: CRN, crambin; BPTI, bovine pancreatic trypsin inhibitor; SOD, superoxide dismutase.

The poly(L-alanine) hypothesis requires demonstrating that there exists a minimum on the poly(L-alanine) hypersurface closely mimicking a given protein's native structure. In *Section 3*, where we consider supersecondary and tertiary structure of long polypeptides, we will almost always have a crystal structure with which to compare the native and the poly(L-alanine) structures. For the secondary structures of small peptides considered in *Section 3*, we rely on the adequacy of the all-atom gas phase potential of blocked poly(L-alanine) since high-level *ab initio* studies (7) indicate that the all-atom model of alanine dipeptide performs quite well structurally and energetically.

Optimization Techniques. We have used two types of optimization procedures for finding minima on the potential energy surfaces described above. For the cases of small peptides (number of amino acids ≤ 10) in secondary structure conformations, we have used a quasi-Newton sequential quadratic programming algorithm (8) with nonlinear constraints. In these cases, five types of constraints were constructed: an L-chirality constraint, a peptide torsion constraint (trans or cis), ϕ (C_{i-1}-N_i-$C_{\alpha i}$-C_i) and Ψ (N_i-$C_{\alpha i}$-C_i-N_{i+1}) constraints appropriate for the secondary structure under consideration, and hydrogen bond constraints, again when appropriate for the secondary structure type. The minimization was started with a structure closely resembling the secondary structure under consideration and minimized with the objective function and the five types of constraints discussed above. The nonlinear constraints were imposed with upper and lower bounds. The optimization was considered converged when the L-chirality dihedral ($C_{\alpha i}$-N_i-C_i-$C_{\beta i}$) constraint of 33° was satisfied to within ±3°, the peptide torsion ($C_{\alpha i}$-C_i-N_{i+1}-$C_{\alpha i+1}$) constraint of 0° or 180° satisfied to within ±10°, the backbone dihedrals ϕ and Ψ converging to within ±0.5° of their secondary structure value, and the hydrogen bond (O_i–H_j) constraint of 1.9 Å falling within a window of ±0.4 Å. The second stage of the procedure is to minimize with no constraints starting with the structure obtained from the constrained first stage.

For the cases of polypeptides and proteins of >10 residues, we have imposed a penalty function protocol for relaxing starting structures into a nearby local minimum. Using a good approximation to the target structure class (secondary, supersecondary, and tertiary) the procedure begins by placing a harmonic penalty function on all heavy atoms,

$$V_p = \sum_i k_p(r_i - r_{io})^2 \qquad [1]$$

and minimizing using the Powell algorithm (9) on the hypersurface defined by Eq. **1** and ref. 6. After a rms derivative convergence of 0.1 kcal/mol·Å (1 cal = 4.18 J) is reached (or after the completion of 200 minimization steps), the penalty function force constant, k_p, and the equilibrium value, r_{io}, are updated by reducing k by 5 kcal/mol·Å^2 and reassigning r_{io} to be the position of i at the completion of the last minimization cycle. Once the penalty function is totally eliminated, the structure is minimized using adopted basis Newton Raphson (6) until the rms derivative is <0.005 kcal/mol·Å.

Section 3: Results

We have found poly(L-alanine) minima corresponding to a wide variety of secondary structure (10–13): α-helix and antiparallel β-sheet conformers for n = 8, 26, 46, 58, 124, and 152; polyglycine II and polyproline for n = 8; and turns of type I, V, and V′ (10) for n = 3 and 8. The nonlinear optimization protocol with constraints (8) outlined in *Section 2* was used to find the secondary structure minima of Ac-$(Ala)_n$-NHMe, for n = 3 and 8, discussed below. The penalty function protocol with the extended-atom potential function presented in *Section 2* was used to find the secondary structure minimum for the larger polypeptides.

We have also found poly(L-alanine) minima corresponding to at least two types of supersecondary structures: the parallel β-sheet–α-helix–parallel β-sheet (β–α–β) and α-helix–turn–α-helix–turn–α-helix (α–t–α–t–α). For the case of α–t–α–t–α, the starting configuration of 30 residues was built "by hand," and energy minimized to give this supersecondary structure type. The initial configuration of the β–α–β secondary structure was taken from a subsequence (residues 200–249) of the crystal structure of carboxypeptidase A (14). Nonaliphatic heavy atom centers were given hydrogens so that all hydrogen positions satisfied excluded volume and geometric constraints (6). (Nonpolar hydrogens are represented by an extended version of the aliphatic carbon to which they are attached.) The resulting hydrogenated crystal structure was minimized using the penalty function protocol presented in *Section 2*. This structure was edited so that the backbone atoms and β-carbon side-chain atoms were retained; for proline and glycine, the backbone hydrogen atom and β-carbon side chain were added, respectively, with all geometric and steric constraints satisfied. This poly(L-alanine) starting structure was minimized again with the penalty function protocol (*Section 2*). The rms deviation between the native sequence and alanine sequence appears in Table 1.

While it is well appreciated that poly(L-alanine) is capable of adopting many different secondary structures, it has not been demonstrated that poly(L-alanine) can mimic wide varieties of tertiary structure. We now demonstrate that this is so for the following representative examples: the 46-residue protein crambin (CRN) (15), the 58-residue protein bovine pancreatic trypsin inhibitor (BPTI) (16), the 124-residue protein A (RNase) (17), and the 158-residue protein superoxide dismutase (SOD) (18).

All crystal structures were taken from the Protein Data Bank (Brookhaven National Laboratory) and supplied with polar hydrogens. The resulting structure was minimized using the penalty function protocol described in *Section 2* and the CHARMM potential with parameters appropriate for all 20 commonly occurring amino acids and terminating groups (6). The resulting four converged structures are the tertiary structure controls (unperturbed by explicit crystal forces) for the comparisons made below. The minimized, hydrogenated native structures were then edited so that only the backbone atoms and the side-chain β carbon atoms were retained. For the cases of proline and glycine, a hydrogen and methyl group were added, respectively, with all geometric and excluded volume constraints satisfied to make L-alanine. This starting poly(L-alanine) structure was then minimized again using the penalty function procedure outlined in *Section 2* and the CHARMM potential with parameters appropriate for the alanine residue (6).

The rms differences between the crystal structure backbone and β carbons and those of the converged, folded poly(L-alanine) counterparts for the four proteins are given in Table 1. There is a remarkable consistency of 2.1–2.6 Å rms

Table 1. rms difference between crystal structure and the poly(L-alanine) analogue

	Difference, Å				
Sequence	Ala	Ala-Cys	Ala-Gly	Ala-Pro	AGP or AGPC
β–α–β	2.07	—	—	—	—
CRN	2.42	2.11	2.75	2.35	1.79
BPTI	2.29	1.80	1.73	2.14	1.53
RNase	2.59	2.31	2.85	2.52	2.62
SOD	2.37	2.43	2.72	2.30	2.17

AGP, Ala-Gly-Pro; AGPC, Ala-Gly-Pro-Cys.

difference between the alanine analogue of the four proteins and the proteins themselves; in fact, the rms differences are similar in magnitude to the resolution of the x-ray diffraction experiments that determined the four native structures examined here. Most of the differences between the native structures and their poly(L-alanine) images arise from small local deformations that are largely pattern preserving; i.e., most important secondary and tertiary fingerprints are retained in the poly(L-alanine) analogue structures. Ribbon structure comparisons of CRN, BPTI, RNase, and SOD vs. their poly(L-alanine) analogues are shown in Figs. 2–5. Both the rms differences and visual comparisons demonstrate the validity of the poly(L-alanine) hypothesis—namely, that there is a short-distance mapping of minima between the homosequence and the native heterosequence for which strong agreement is obtained along the backbone and at β-carbon side-chain positions. In addition, absolute energies for poly(L-alanine) in the extended conformer, the α-helix, and the native-like conformation of the four proteins are presented in Table 2. Except for the CRN case, the native-like structure is a high-energy minimum on the poly(L-alanine) surface; the α-helix conformer is most likely the ground state.

The proteins CRN and BPTI are cases that have nontrivial tertiary structure in addition to secondary structure but that are not large enough to exhibit a sizeable and tightly packed hydrophobic core. As shown in Fig. 2, the CRN native structure corresponds to the following sequence of secondary structures: coil 1 and helix 1 (residues 1–19), loop, helix 2, coil 2, and turn (residues 20–44), and coil 3 (residues 44–46). Two disulfide bonds link helices 1 and 2 at the loop end and the coil 2 end, and a third disulfide links coil 2 and coil 3. In the case of the poly(L-alanine) version of CRN, the ordering of secondary structure is largely the same except that the large loop is now a tight turn between the helices and the second helical segment commences sooner than it does in native CRN. The rms differences between CRN and its poly(L-alanine) analogue is 2.42 Å. A description of the secondary structure for BPTI is coil for residues 1–13; two turns around residues 14 and 38, respectively; β-sheet for amino acids 16–25 and 28–36; and helix for residues 47–58. Disulfide bonds link residues 5 and 55, 14 and 38, and 30 and

FIG. 2. Backbone, ribbon structure comparison of CRN and the CRN native-like conformer of Ala_{46}. (*Upper*) Native CRN backbone. (*Lower*) Ala_{46} backbone mimicking native CRN.

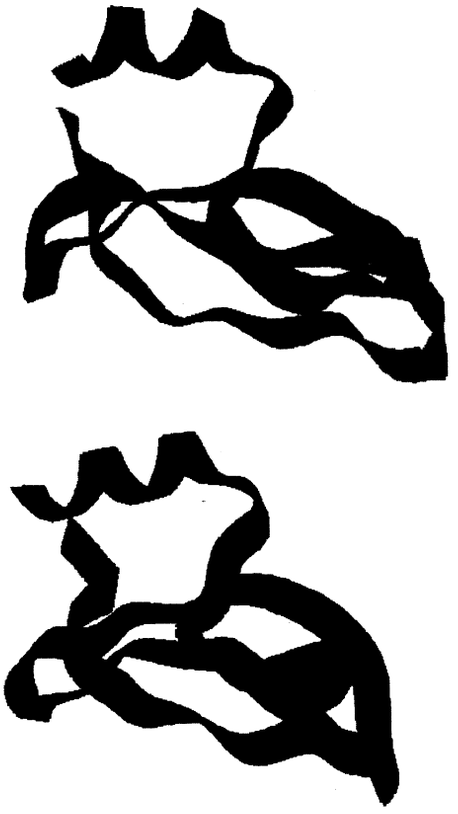

FIG. 3. Backbone, ribbon structure comparison of BPTI and the BPTI native-like conformer of Ala_{58}. (*Upper*) Native BPTI backbone. (*Lower*) Ala_{58} backbone mimicking native BPTI.

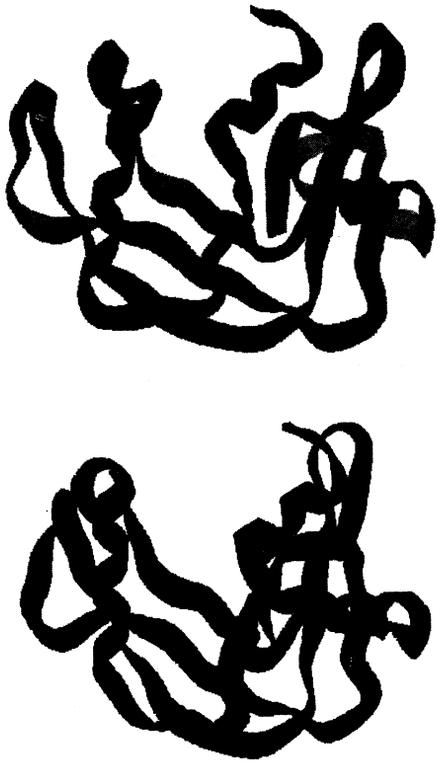

FIG. 4. Backbone, ribbon structure comparison of RNase A and the RNase A native-like conformer of Ala_{124}. (*Upper*) Native RNase A backbone. (*Lower*) Ala_{124} backbone mimicking native RNase A.

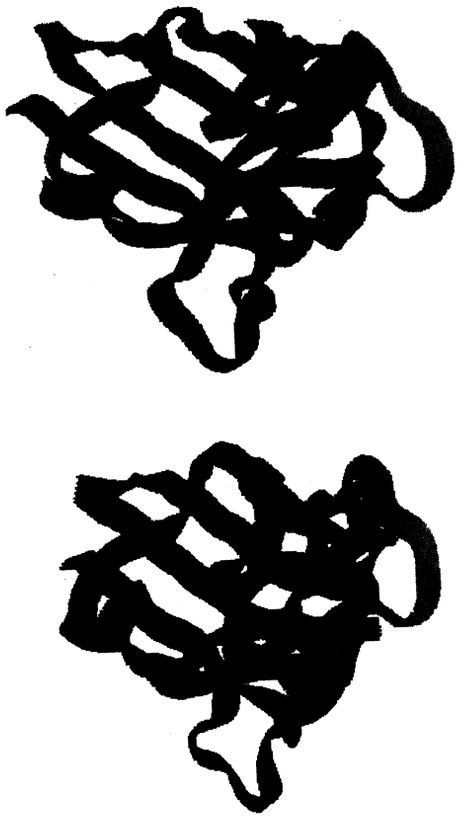

FIG. 5. Backbone, ribbon structure comparison of SOD and the SOD native-like conformer of Ala_{152}. (*Upper*) Native SOD backbone. (*Lower*) Ala_{152} backbone mimicking native SOD.

51. The poly(L-alanine) analogue shows a similar ordering of secondary structure, and the overall rms difference between the native structure and the poly(L-alanine) analogue is 2.29 Å.

Both RNase and SOD are rather large globular proteins that contain tightly packed cores, but they differ greatly in their predominant secondary structure: RNase is largely helical while SOD is an example of a β-barrel. RNase and SOD show rms differences with their poly(L-alanine) analogues of 2.59 Å and 2.31 Å, respectively. As we will show in *Section 4* some evidence indicates that the lack of side-chain electrostatic interactions in $(L\text{-Ala})_M$ is largely responsible for the observed rms difference between the native sequence and the poly(L-alanine) structures and that little collapse has resulted due to the elimination of large residues in the core.

Section 4: Poly(L-Alanine) as a Reference State

Support for the poly(L-alanine) hypothesis adduced in *Sections 1–3* provides evidence that the L-alanine homosequence may be a universally good reference state for understanding the structures and energies of real proteins. In fact a statistical mechanical perturbation treatment can be derived, based on the poly(L-alanine) reference state, as a means for understanding the free energy stabilization of native protein tertiary structure minima. Two possible kinds of questions can be addressed with such an approach: (*i*) the relative importance of packing and electrostatics as one perturbs alanine into larger side chains and/or polar residues and (*ii*) the degree to which the *specific* native sequence is required for obtaining the optimal tertiary structure.

Table 2. Relative energies of poly(L-alanine) minima

Sequence	Residues, no.	Relative energy, kcal/mol		
		Tertiary structure	α-Helix	Extended
CRN	46	−1339.34	−1335.86	−1054.47
BPTI	58	−1642.35	−1699.43	−1326.44
RNase	124	−3340.17	−3697.10	−2820.99
SOD	152	−4119.16	−4531.84	−3423.95

As an initial foray we have investigated these questions by performing a penalty function minimization (*Section 2*) on heterosequences of four proteins—CRN, BPTI, RNase A, and SOD—where either (*i*) all cysteines were retained when disulfide linkages are present in the native state, (*ii*) all amino acids were changed to alanine except for the proline positions, (*iii*) all residues were changed to alanine except for the glycine positions, or (*iv*) the sequence was mutated to all alanine except for glycine, alanine, and cysteine, which were retained from the native sequence. We have distinguished proline and glycine from the remaining 19 amino acids because their local conformation space should be most different from that of alanine (and the other residues). The cysteines are also unusual because they can form covalent disulfide bonds nonlocal in sequence. The minimized, hydrogenated crystal structures were edited so that one of the above subset of amino acids was retained, while the remaining amino acids were converted into L-alanine (*Section 3*). This starting poly(L-alanine) copolymer was then minimized in energy using the penalty function procedure outlined in *Section 2*; the resulting rms differences between the native structure and the four types of subset amino acids in their original sequence positions in the poly(L-alanine) sequence are given in Table 1.

We have found that evaluating "local" rms differences, where local refers to smaller portions of the sequence, provides a means for comparing the structural differences between each protein and its poly(L-alanine) analogue. In the case of CRN, the overall rms difference between the native sequence and poly(L-alanine) structure is 2.42 Å. The rms differences for the subset of amino acids 1–10, 11–20, 21–30, 31–40, and 41–46 are 1.1 Å, 1.1 Å, 1.8 Å, 1.8 Å, and 1.2 Å, respectively. These local rms differences indicate that the structure is largely sequence independent in the first half of the chain but that either a sequence dependence exists, or only a few amino acids are important, in the latter half of the CRN structure. We find that reintroducing the cysteines into their native sequence positions in poly(L-alanine) in order to form three covalent disulfide links results in an overall rms improvement of 0.3 Å and in improvement in each of the subsequences as well. The disulfide bonds pull portions of the sequence together to make the structure more rigid; the poly(L-alanine) structure is a little more "floppy" by comparison. Glycine and proline do not provide a global improvement in the rms differences (2.75 and 2.35 Å, respectively), but they do provide local improvement in the latter half of the CRN structure. The reintroduction of proline-22 improves the local rms for the subsequence residues 21–30 by 0.6 Å, whereas glycine-37, which participates in the turn, improves the local rms for the subsequence residues 31–40 by 0.6 Å.

For BPTI, the rms difference between the native structure and the poly(L-alanine) analogue of 2.29 Å can be improved globally (1.80 Å) by reincorporating the cysteines, whose cross linkages are important for rigidifying the protein. The BPTI subsequences 1–10, 11–20, 21–30, 31–40, 41–50, and 51–58 have rms differences of 1.6 Å, 1.6 Å, 1.4 Å, 1.3 Å, 1.4 Å, and 1.6 Å, respectively. Local improvement in regions 11–20 and 31–40 is provided by reintroducing glycine, whose larger local conformation space allows a better description of the turn regions and results in a global improvement of 0.6 Å. This is consistent with past studies using simplified protein representations, where glycine is known to play a crucial structural role in the turn regions of native protein conformations (19).

For the globular proteins RNase and SOD, we have found that the radius of gyration of the poly(L-alanine) analogues has decreased by only 1 Å when compared to the native sequence conformation. This indicates that the loss of large side chains does not result in collapse of the poly(L-alanine) structure to retain a compact nonpolar core. In addition, the substitution of cysteine, glycine, and proline into the poly(L-alanine) matrix does not result in any significant improvement in the rms values of 2.59 Å for RNase and 2.37 Å for SOD. Further examination of the structural differences between the native sequence and the poly(L-alanine) structures indicates that hydrogen bonding interactions between side chains may account for the changes in rms differences for these two proteins and that the strong structural influences of cysteine and glycine found for small proteins may be of only secondary importance in globular proteins. Although these are preliminary results for two isolated cases, they indicate that a systematically applied perturbative approach may uncover some general principles regarding the relative importance of packing, electrostatic interactions, and the influence of special amino acids such as proline, cysteine, and glycine.

Section 5: Discussion and Conclusions

This work provides validation of the poly(L-alanine) hypothesis: namely, that there exists a short-distance mapping between the native structure minimum of any given polypeptide hypersurface and a local minimum on the poly(L-alanine) hypersurface for which the backbone and β-carbon structural agreement is very close. One significant implication is that poly(L-alanine) may serve as a universally good reference state in a perturbative treatment designed to evaluate free energy driving forces and to understand the importance of particular amino acids that stabilize the folded state of the native sequence.

Recent work by Matthews and coworkers (2) has explored the use of alanine as a "generic" residue to replace "nonessential" residues in T4 lysozyme, to see whether the resulting mutant is correctly folded and functional. The mutant T4 lysozyme contained alanine substitutions at positions 128, 131, 132, and/or 133, so that the number of introduced alanines is a very small fraction of the total number of amino acids; in all cases except one, the mutant form was more stable than wild-type lysozyme (2). More recently the same group has substituted alanine for native residues within the hydrophobic core (3) and observed that such mutations are "cavity creating"; i.e., proteins do not collapse to avoid vacant space within the core (8). Our model poly(L-alanine) studies also show such an effect in the case of the larger proteins, RNase A and SOD, where the radius of gyration was found to decrease by only 1.0 Å upon *full* mutation.

The present results and the mutagenesis experiments of Matthews and coworkers (2, 3) indicate that not all details of the native sequence are important for accessing a compact conformation closely resembling the given native state. In *Section 3* we have shown that the "generic" homopeptide poly(L-alanine) is capable of folding into a good likeness of a variety of native structure proteins. Thus another important implication of the poly(L-alanine) hypothesis is that the structure predictor problem in the protein folding problem may not require the specific native sequence for mapping onto a good structural likeness of the native minimum—i.e., that specific amino acids may be important only for lowering the free energy. This line of thought has been pursued by Hagler and Honig (19) and by Levitt and Warshel (20) using simplified representations to fold protein native states. Energy minimization studies of a one-center side-chain representation of BPTI (20) did not predict important topological features of the BPTI fold and exhibited an overall rms error of 6.0 Å. Our studies indicate that minima can be found on a "simplified" surface whose structural likeness to the protein in question is very good; the poly(L-alanine) analogue of BPTI, for example, reproduces the topological features of loop threading and the 180° twist in the β-sheet missed by past folding studies (19). Whether simplified representations provide the correct pathway to this minimum is another question altogether; the studies reported in refs. 19 and 20 seem to indicate that they do not.

1. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1973) *J. Mol. Biol.* **7**, 95–99.
2. Zhang, X.-J., Baase, W. A. & Matthews, B. W. (1991) *Biochemistry* **30**, 2012–2017.
3. Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992) *Science* **255**, 178–183.
4. Head-Gordon, T., Stillinger, F. H. & Arrecis, J. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11076–11080.
5. Momany, F. A., Klimkowski, V. J. & Schäfer, L. (1990) *J. Comp. Chem.* **11**, 654–662.
6. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comp. Chem.* **4**, 187–217.
7. Head-Gordon, T., Head-Gordon, M., Frisch, M. J., Brooks, C. L. & Pople, J. A. (1991) *J. Am. Chem. Soc.* **113**, 5989–5997.
8. Gill, P. E., Murray, W., Saunders, M. A. & Wright, M. H. (1986) *User's Guide for* NPSOL, a Fortran package for nonlinear programming, Report SOL 86-2 (Dept. of Operations Res., Stanford University, Stanford, CA).
9. Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, V. T. (1986) *Numerical Recipes* (Cambridge Univ. Press, Cambridge, U.K.).
10. Creighton, T. E. (1984) *Proteins, Structures and Molecular Properties* (Freeman, New York).
11. Asakura, T. (1981) *Makromol. Chem.* **182**, 1153–1165.
12. Tobias, D. J., Sneddon, S. F. & Brooks, C. L. (1991) *Adv. Biomol. Simul.*, in press.
13. Tobias, D. J., Sneddon, S. F. & Brooks, C. L. (1991) *J. Mol. Biol.* **216**, 783–796.
14. Rees, D. C., Lewis, M. & Lipscomb, W. N. (1983) *J. Mol. Biol.* **168**, 367–387.
15. Hendrickson, W. A. & Teeter, M. M. (1981) *Nature (London)* **290**, 107–113.
16. Deisenhofer, J. & Steigemann, W. (1975) *Acta Crystallogr. Sect. B. Struct. Crystallogr. Cryst. Chem.* **31**, 238–250.
17. Kartha, G., Bello, J. & Harker, D. (1967) *Nature (London)* **213**, 862–865.
18. Tainer, J. A., Getzoff, E. D., Beem, K. M., Richardson, J. S. & Richardson, D. C. (1982) *J. Mol. Biol.* **160**, 181–217.
19. Hagler, A. T. & Honig, B. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 554–558.
20. Levitt, M. & Warshel, A. (1975) *Nature (London)* **253**, 694–698.